CORONAVIRUS

Interdependent program evaluation: Geographic and social spillovers in COVID-19 closures and reopenings in the United States

Michael Zhao[1], David Holtz[2,1], Sinan Aral[1,3]*

Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

In an interconnected world, understanding policy spillovers is essential. We propose a program evaluation framework to measure policy spillover effects and apply that framework to study the governmental responses to COVID-19 in the United States. Our analysis suggests the presence of social spillovers. We estimate that while state closures directly reduced mobility by 3 to 4%, all other states locking down further decreased mobility in the focal state by 8 to 14%. Similarly, while reopening directly increased mobility by 2 to 3%, all other states' reopening increased mobility in the focal state by 12 to 21%. Our analysis also suggests geographic spillovers: Travel from locked down origins to open destinations increased by 12 to 29%. In contrast, travel from reopened origins to locked down destinations decreased by 6 to 7% for nearby counties and by 14 to 18% for distant counties. Despite its limitations, we believe that our approach takes the first steps toward creating a framework for interdependent program evaluation across policy domains.

INTRODUCTION

Policy-makers and academics often consider the efficacy of policy interventions in isolation; however, different jurisdictions are connected both geographically and, increasingly, socially. This interconnectedness means that the policies enacted by one governing body affect outcomes beyond the area in which the policy is expected to have a direct effect (*1*). For instance, previous studies have identified the presence of both geographic and social spillovers in contexts such as voting, crime reduction, and physical exercise (*2*–*4*). As the world continues to grow more interconnected, program evaluation methods need to be systematically expanded to assess the effectiveness of policies that exhibit meaningful spillover effects.

Government policy responses to coronavirus disease 2019 (COVID-19) provide a recent and relevant example of the type of policy intervention in which the interdependence between different jurisdictions is important. In the early months of the pandemic, countries around the world imposed strict limits on human mobility by banning gatherings, closing down nonessential businesses, and instituting shelter-in-place orders. Although human mobility levels were already decreasing before the enactment of such policies (*5*, *6*), these measures further reduced the spread of the virus (*5*–*9*) and led to lower morbidity and mortality rates. These policies had effects beyond the counties or states in which they were implemented. These spillover effects need to be explicitly considered, as they can have profound policy implications. For instance, Holtz *et al.* (*10*) found that cross-county spillovers from shelter-in-place policies exceeded the direct impacts of local policies, underscoring the importance of governmental coordination to reduce a potential "loss from anarchy" due to piecemeal implementation of closure policies across regions.

Following the lockdowns, many countries began reopening, allowing the coronavirus room to spread. For example, after reopenings began in the United States, several COVID-19 hot spots emerged, causing some local governments to reimpose social distancing measures. As different regions of the world continue to relax and tighten mobility restrictions, it is critical that we understand how different mobility policy regimes contribute to increases and decreases in COVID-19 case counts. While there is abundant research on the efficacy of shutting down, the body of quantitative research examining the impact of reopening and/or the factors that make reopening safe and effective is still nascent (*11*–*15*). Although most of the research is not causal (*13*, *15*), the more causally rigorous studies (*14*) neglect the magnitude of a policy's spillover effects relative to its direct effects and do not account for non–travel-related social spillovers. If reopenings cause substantial increases in mobility and exhibit strong spillover effects, then countries that reopen without national coordination could face substantial difficulty in controlling the resurgent spread of the coronavirus and future pandemics. Furthermore, understanding whether the effects of closures and reopenings are symmetric or asymmetric is important in determining the optimal adaptive policy in response to the ebbs and flows of local COVID-19 case rates.

In this work, we combine data from a variety of different sources including the mobility data of more than 22 million mobile devices, daily data on state-level closure and reopening policies, social media connections among 220 million U.S. Facebook users, temperature and precipitation data from 62,000 weather stations, county-level census data, county-level COVID-19 case and death counts, and county- and state-level unemployment. We use these data to measure the direct causal impacts of a state's COVID-19 closure and reopening policies on its own mobility patterns; the causal spillover effects of other, socially connected states' closure and reopening policies on a state's mobility patterns; and the impacts of both origin- and destination-county closure and reopening policies on cross-county mobility, capturing the travel-related spillover effects created by uncoordinated policies implemented across different states and counties. In doing so, we contribute to a growing literature that uses population-scale digital trace data (*16*) to study the impacts of mobility-related interventions in response to COVID-19 and what makes them successful (*5*, *17*). Researchers have shown, for example, that demographic attributes (*18*), political partisanship (*19*, *20*), broadband access (*21*), belief in science (*22*), and information exposure (*23*, *24*) moderate compliance with social distancing policies.

[1]MIT Initiative on the Digital Economy, 100 Main St., Cambridge, MA 02142, USA. [2]Haas School of Business, University of California, Berkeley, 2220 Piedmont Ave., Berkeley, CA 94720, USA. [3]MIT Sloan School of Management, 100 Main St., Cambridge, MA 02142, USA.
*Corresponding author. Email: sinan@mit.edu.

Downloaded from https://www.science.org on July 24, 2022

Our empirical methodology is grounded in a reduced form econometric approach called difference-in-differences (DiD), which is widely used across economics, political science, and public health for policy evaluation. One of the key identifying assumptions of DiD is the "parallel trends" assumption, which states that in the absence of the policy being studied, the time series trends for units that enact the policy (treated units) and those that do not (untreated units) would move in parallel. This assumption is often tested by verifying that the treated and untreated lines move in parallel before the policy intervention in question. Figure 1A shows the time series trends for the four human mobility measures we track, each of which is constructed using data provided by SafeGraph: the daily average number of locations visited by mobile devices, the proportion of devices traveling more than 2 km, the proportion of devices that spend over an hour away from home, and the proportion of devices leaving home. The trend lines for each outcome suggest that in our context, the parallel trends assumption is credible. For DiD to yield unbiased causal estimates, the timing of the treatment also needs to be (conditionally) exogenous. To, as well as is possible, account for potential endogeneity in the timing with which different states enacted different policy interventions and also to make our DiD models robust to deviations from the parallel trends assumption, our DiD models control for weather, COVID-19 case counts, COVID-19 deaths, and unemployment levels using daily temperature and precipitation data from the Global Historical Climatology Network (*25*), daily COVID-19 case and death counts from the *New York Times*, and unemployment data from the U.S. Bureau of Labor Statistics.

Our state-level data on closure and reopening policies come from the COVID-19 U.S. State Policy Database (*26*). During our observation period, different U.S. states implemented a number of different policies, including bans on large gatherings; closures of businesses such as gyms, movie theaters, and restaurants; and shelter-in-place orders with varying degrees of strictness. To create sufficient statistical power to identify causal effects, we collapsed the many different policy interventions into three general classes of policies based on their level of restrictiveness and identify the "periods" during which different states are subject to these policy types. The policy types we analyzed are as follows: "initial policies" (ip), which covers the set of moderately restrictive policies from when a state implements its first closure policy of any kind until it implements a shelter-in-place order; "shelter-in-place" (sh), which covers the more restrictive statewide shelter-in-place orders or stay-at-home orders; and "reopening" (ro),

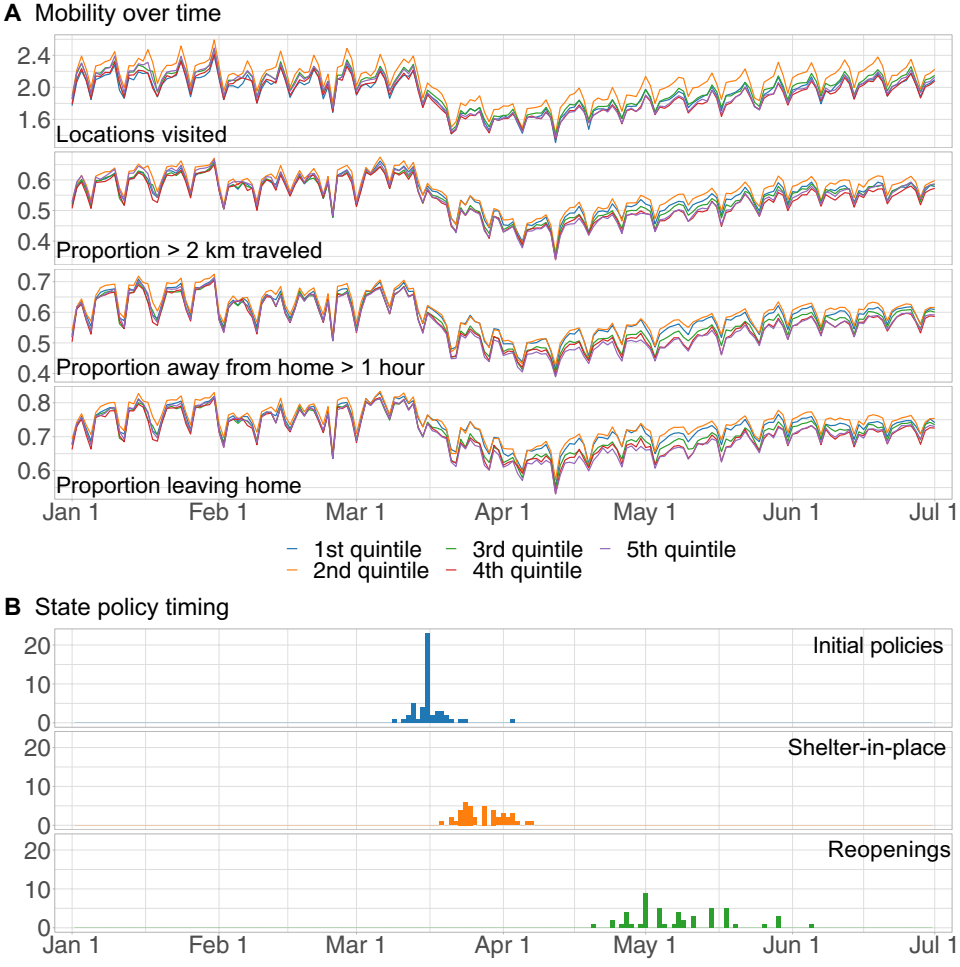

Fig. 1. Data on mobility behaviors and state policy timing. (**A**) The time series trends for the number of locations visited per device, the proportion of devices traveling more than 2 km, the proportion of devices spending more than an hour away from home, and the proportion of devices leaving home. Each color represents the averages across clusters of 10 states grouped by how soon they reopened. (**B**) The count of states that enter into a particular policy period on each day.

Downloaded from https://www.science.org on July 24, 2022

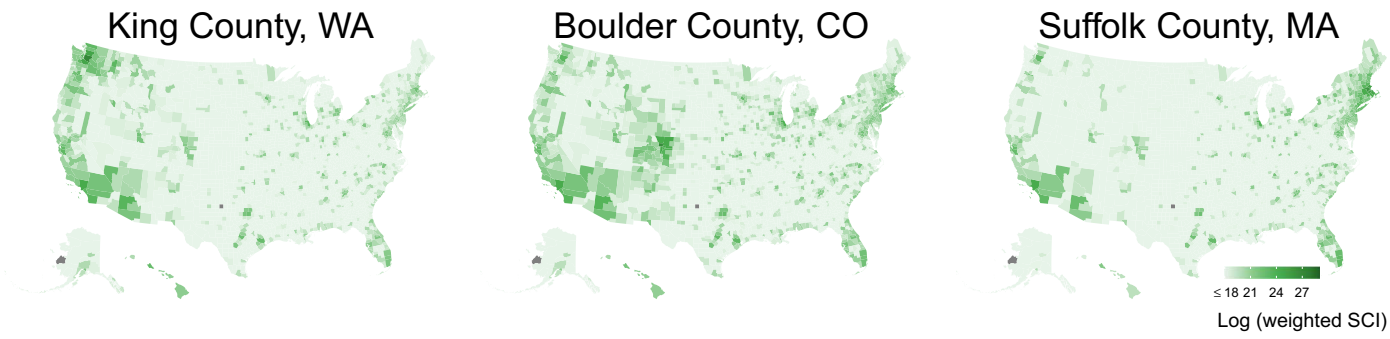

Fig. 2. Data on social connectedness. This figure shows examples of the population-weighted SCI used to construct socially weighted measures of alter state policies and behavior.

which covers various reopening plans. We also replicated our main results using an alternative middle policy period definition that accounts for a broader set of "movement restriction" policies, as opposed to just shelter-in-place and "stay-at-home" orders (see section S4.3). In all of our analyses, we interpret the effect of the initial policies as a weighted sum of the causal effects of many different mobility-curbing interventions implemented by states across the United States at different points in time, the effect of shelter-in-place as the causal effect of implementing some type of shelter-in-place order, and the effect of reopening as the causal effect of enacting some type of reopening plan. It is worth noting that collapsing COVID-19–related mobility policies into these three broad classes prevents us from measuring the causal effect of any specific policy (e.g., banning indoor dining at restaurants) but enables our causal estimates of direct and indirect policy effects nationwide.

The timing with which different U.S. states entered different policy periods is shown in Fig. 1B. Many states implemented their initial policies in mid-March, during which all four mobility measures were falling (see Fig. 1A). Mobility continued to fall during late March and early April, during which most states implemented some sort of shelter-in-place order. States began enacting reopening plans in late April and continued doing so through early June. During this period, mobility levels rose gradually until they returned to the same levels observed before March.

In addition to measuring the direct effects of a given state's policies on its own mobility levels, we also measured the effect of other, geographically and socially "connected" states' policies on each state's mobility levels. Throughout the paper, we adopt language from the networks literature and refer to the state whose mobility levels we measure as the "ego" state and refer to the states to which it is connected as that state's "alter" states. To construct the weighted set of counties to which a given county is connected, we combined Facebook's "Social Connectedness Index" (SCI) (*27*), which uses an anonymized snapshot of the Facebook social network to provide a measure of the intensity of social connectedness between geographic locations, with county-level census population estimates to create a population-weighted SCI. The larger the value of this index for a given ego-alter county pair, the larger the fraction of ego county Facebook friendships that are with people in the alter county. More details on the construction of the population-weighted SCI can be found in sections S1.3 and S1.4, and verification that our results are robust to alternative transformations of the population-weighted SCI can be found in section S4.2. Figure 2 shows choropleth map visualizations of the population-weighted SCI for three different ego counties (King County, WA; Boulder County, CO; and Suffolk County, MA) and each possible alter county. For a given ego county, the alter counties with the highest connectedness index are typically a combination of nearby counties and far away counties with large populations.

We also used DiD models to measure the effect of different policies and policy contrasts on cross-state mobility patterns, e.g., how is travel from one state to another affected when one state has begun reopening, whereas the other is still subject to a shelter-in-place order? Figure 3 provides some insight into how policy contrasts can affect cross-county and cross-state mobility. In the weeks following San Francisco County's lockdown, travel into the county from other counties decreased, whereas in the weeks following the same county's reopening, inbound travel from other counties increased. The same pattern is also observed for Chatham County, GA.

RESULTS

We first estimated a "no spillovers" DiD model that accounted for the ego state's mobility-related policies, but not the policies of the ego state's alter states. The results of that model are found in the top row of Fig. 4. The leftmost column shows that according to the no spillovers model, the combination of initial mobility-related policies deployed across the United States did not cause a statistically significant reduction in mobility, as measured by our four mobility outcomes. In contrast, the middle column of Fig. 4 shows that according to the no spillovers model, statewide shelter-in-place orders reduced mobility within a state by 2.8 to 3.5% on average, depending on the mobility outcome used. Last, the rightmost column shows that according to the no spillovers model, mobility increased by 2.0 to 2.8% on average once a state reopened, depending on the mobility outcome used.

We next estimated a "spillovers" DiD model, in which we accounted not only for the ego state's policy but also the policies of the ego state's alter states. In this model, our estimates of the direct effects of each type of policy on human mobility remained quantitatively and qualitatively consistent with the estimates obtained from the no spillovers model (see the top row of Fig. 4). Our estimates of the spillover effects caused by different mobility-related policies are found in the bottom row of Fig. 4. The spillovers model suggests the existence of spillovers caused by social distancing policies, shelter-in-place orders, and reopenings. The leftmost column shows that according to the spillovers model, when the ego county's alter states began implementing initial social distancing policies, ego county mobility dropped by 6.1 to 8.2%, depending on the mobility outcome used. Furthermore, the middle column shows that according to the spillovers model, when

Downloaded from https://www.science.org on July 24, 2022

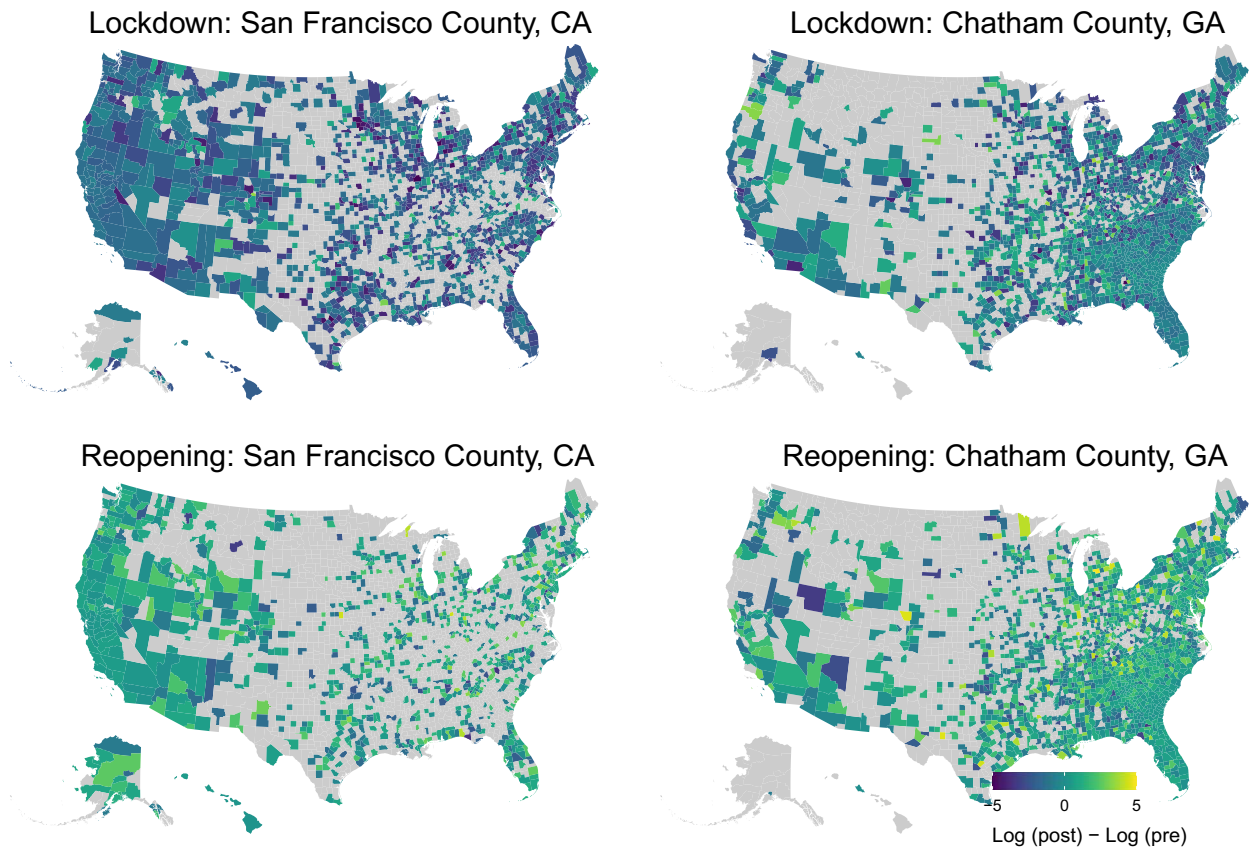

Fig. 3. Changes in cross-state mobility patterns after policy changes. This figure shows examples of the difference in travel to a destination county for the 3 weeks before and after a lockdown or reopening.

an ego county's alter states implemented shelter-in-place, ego county mobility dropped by another 8.3 to 11.7%. Last, the rightmost column shows that according to the spillovers model, when an ego county's alter states began reopening, ego county mobility increased by 12.1 to 21.4%. All of these effects account for the ego counties' policies and estimate the additional effects of alter states' policies on ego mobility.

Figure 5 shows the results from a cross-state travel DiD model that measured the effect of origin- and destination-county policies on travel from the origin county to the destination county. The top row shows the effects on the number of devices traveling from the origin county to the destination county, whereas the bottom row shows the effects on the proportion of devices traveling from the origin county to the destination county. Overall, these results show evidence plausibly consistent with the existence of travel-driven policy spillovers. The right column shows the effects of the destination county implementing different policies. We found that when a destination county was subject to a statewide shelter-in-place order, it received 2.3 to 6.8% less cross-state traffic, depending on the outcome metric used (see blue points). In contrast, being subject to a statewide reopening plan increased cross-state travel to destination counties by 4.2 to 5.4%. In addition to communicating overall effects (blue points), Fig. 5 also communicates the heterogeneous effects of origin- and destination-county policies for distant pairs of counties (orange points) and nearby pairs of counties (green points). We observe notable levels of heterogeneity; for instance, we estimate that destination counties in distant pairs being subject to statewide reopenings increased cross-state travel by 4.5 to 5.3%, whereas destination counties in nearby pairs being subject to statewide reopening plans increased cross-state travel by 8.2 to 9.9%. Insofar as we observe differences in the effects of origin and destination policies across our two outcome variables, we believe that this is due to the fact that the proportion of origin devices traveling to the destination county is bounded between zero and one, whereas the number of devices traveling from the origin county to the destination county is much more highly dispersed. Counties in which there are many origin county devices therefore have much more influence on our "number of devices" estimates.

Last, when we expanded our cross-state travel DiD model to include all possible origin-destination policy interactions (Fig. 6), several additional findings emerged. The top row of Fig. 6, for example, shows the effects on the number of devices traveling from the origin county to the destination county, whereas the bottom row shows the effects on the proportion of devices traveling from the origin county to the destination county. The second column from the left shows that when origin counties were subject to their own initial policies, the enactment of destination policies did not appear to measurably affect cross-state mobility. However, we estimate that when origin counties were subject to a shelter-in-place order (second column from the right), travel to distant counties that were subject to initial distancing policies decreased by 5.6 to 11.2% (orange points), whereas travel to nearby counties that were subject to initial

Downloaded from https://www.science.org on July 24, 2022

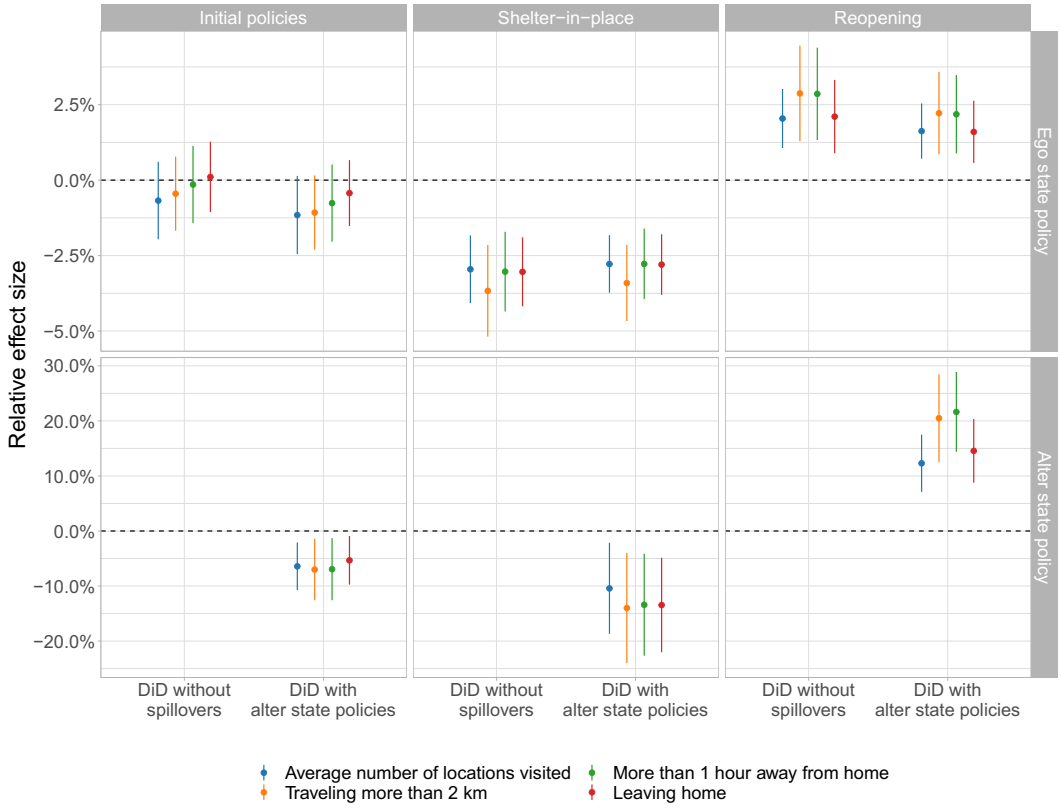

Fig. 4. Empirical estimates of the effects of ego state policy and alter state policy on county-level mobility. This figure depicts the estimated impact of both ego and alter state policies. (**Top**) Corresponds to ego state's policy periods and (**bottom**) corresponds to the alter states' policy periods. The *x* axis denotes the model specification used to generate the estimates. All estimates are produced using weighted least squares (WLS), with weights determined by county population. SEs are clustered at the ego county U.S. state level.

distancing policies increased by 29.0 to 29.6% (green points). Similarly, we estimate that a destination county reopening while an origin county was still subject to a shelter-in-place order caused cross-state travel from the origin county to the destination county to increase by 11.8 to 15.9% for nearby counties and by 4.2 to 5.4% for distant counties. In contrast, the rightmost column of Fig. 6 shows that origin counties reopening while destination counties were still subject to shelter-in-place appear to have decreased travel from the origin county to the destination county by 4.5 to 5.1% for nearby counties and by 7.8 to 12.6% for distant counties. Last, we estimate that both origin and destination counties reopening caused travel from origin counties to destination counties to increase by 6.8 to 7.1% for distant destination counties, whereas it did not have an effect for nearby destination counties.

DISCUSSION

Our analysis highlights the importance of taking into account both geographic and social spillovers when evaluating the efficacy of uncoordinated local policy interventions and takes the first steps toward creating a causal framework for measuring the effects of a given policy not only on outcomes in the jurisdiction in which it is enacted but also on outcomes in the jurisdictions to which it is in some way connected. Our study contributes to an emerging literature on the impact of COVID-19 reopening policies (*13*–*15*) and, to our knowledge, is the first to attempt to causally estimate the magnitude of reopening policy spillovers, to explicitly estimate cross-state travel spillovers due to reopening, and to compare estimates of the causal effects of reopening to estimates of the causal effects of related interventions, such as shelter-in-place orders and other policies aimed at curbing human mobility.

That said, this work has a number of limitations. First, there are limitations associated with the data we used to conduct our analyses. Although the SafeGraph panel is sufficiently large to minimize concerns about sampling error, it may exhibit sampling bias as mobile device ownership varies considerably by age and income (*28*). While SafeGraph has shown that their panel is geographically consistent with the U.S. Census population estimates (*29*), it is not clear whether certain demographics are over- or underrepresented, as no device-level demographic data are collected by SafeGraph. Although it is reassuring that we find similar results when using mobility data provided by Facebook as a robustness check (see section S4.1), there are concerns about the representativeness of Facebook's data as well. It is also possible that mobility datasets from both SafeGraph and Facebook are inaccurate in other, unobservable ways.

Second, there are multiple policy-related questions that are extremely relevant but are beyond the scope of our analysis. For instance, our analysis does not account for the relatively few cases in which local- or county-level policies differed from state-level policies and is restricted to mobility outcomes (i.e., we purposefully avoid extrapolating to health outcomes such as morbidity and mortality rates). We encourage follow-up work on both of these fronts. Accounting for

Downloaded from https://www.science.org on July 24, 2022

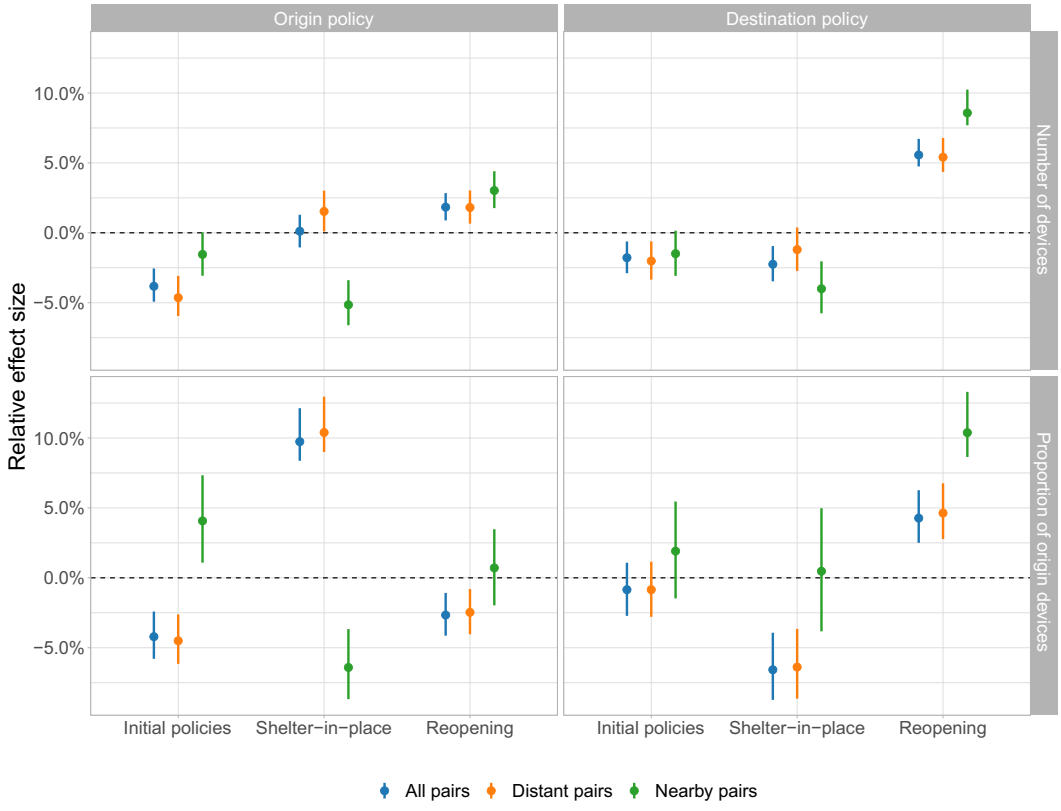

Fig. 5. Empirical estimates of origin and destination policies on cross-state travel. In this figure, (**left**) corresponds to the origin policy periods, while (**right**) corresponds destination policy periods (denoted across the *x* axis as initial policies, shelter-in-place, and reopening). (**Top**) reflects policy effects on the number of devices moving from an origin to a destination, estimated using ordinary least squares (OLS), and (**bottom**) reflects policy effects on the proportion of origin devices moving from an origin to a destination, estimated using WLS with weights proportional to origin county population. Colors correspond to estimates produced using all pairs (blue), distant pairs (>100 km) (orange), and nearby pairs (<100 km) (green). Two-way origin and destination state clustered SEs are used to compute 95% confidence intervals.

disputes between states and localities will likely enable more accurate assessments of the impact of state-level policies. For example, such disputes may have affected the extent to which state-level policy interventions effectively limited mobility during the COVID-19 pandemic. In addition, while it is widely believed that reductions in mobility drive reductions in new infections and their associated deaths (*30*), rigorously establishing the causal chain from cross-state spillovers to infection rates and deaths is beyond the scope of this paper. Our analysis also does not explore the impacts of specific reopening policies (e.g., resuming restaurant dine-in service or lifting gathering restrictions) and instead measures the average causal effects of different "types" of policies, as defined through "policy periods," on human mobility.

Third, our analysis may suffer from model misspecification and/or failure to properly quantify the uncertainty associated with our estimates. With regard to model misspecification, our analyses assume that the relationships between the outcomes and independent variables are all linear. Although estimates produced using a "double machine learning" (*31*) estimator (see section S4.7), which can account for nonlinear relationships in the control variables, were nearly the same as those presented in the paper, this method is designed to uncover heterogeneity in the treatment effect estimates themselves, not to test robustness to model misspecification. With regard to the quantification of uncertainty, although our results remain statistically significant when calculating adjacency- and cluster-robust standard errors (SEs) and when conducting Fisherian randomization inference (see sections S4.5 and S4.6), it is possible that there is still correlation structure in the assignment of our treatment variables that we do not properly account for.

Last, although this work represents the initial step toward creating a causal framework for interdependent program evaluation, unambiguous causal identification is likely beyond our reach given the data at hand. To yield causal estimates, DiD relies on both the parallel trends assumption and the assumption that the timing of the treatment is exogenous conditional on observables. Our investigation of potential anticipatory or lagging behaviors (see section S4.4) suggests that the parallel trends assumption is plausible in our setting. However, conditional exogeneity of treatment timing is more difficult to verify. Although our analyses make use of a robust set of controls, it is possible that both our ego policy "treatment" and our alter policy "treatment" suffer from endogeneity. The fact that our ego state policy estimates do not meaningfully change after the inclusion of alter state policies is encouraging, but our estimates of ego state policy effects may yet be confounded by unobserved, time-varying factors that affect both ego state policy enactment and mobility. For instance, COVID-19–related (mis)information shocks may affect both state-level policy and human mobility. Causal interpretation of our alter state policy coefficients presents other challenges, and it is possible that causally interpreting our alter state policy coefficient estimates is an instance of what epidemiologists refer to as the "table 2

Downloaded from https://www.science.org on July 24, 2022

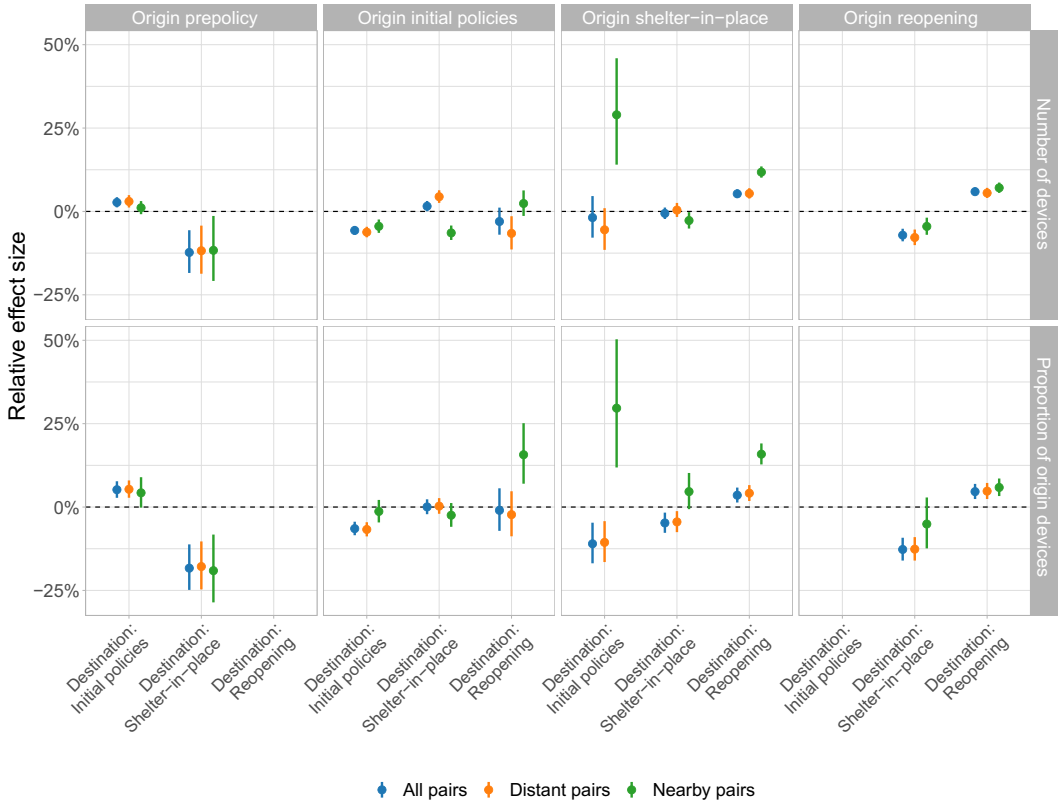

Fig. 6. Empirical estimates of origin and destination policy interactions on cross-state travel. Each block in the figure corresponds to different origin policy periods: prepolicy, initial policies, shelter-in-place, and reopening from left to right. (**Top**) reflects policy effects on the number of devices moving from an origin to a destination, estimated using OLS, and (**bottom**) reflects policy effects on the proportion of origin devices moving from an origin to a destination estimated using WLS with weights proportional to origin county population. Different values along the *x* axis of each column correspond to different destination policy periods: destination initial policies, destination shelter-in-place, and destination reopening. Within each column, the marginal effects of each destination policy given the origin policy. Colors correspond to estimates produced using all pairs in blue, distant pairs (>100 km) in orange, and nearby pairs (<100 km) in green. Two-way origin and destination state clustered SEs are used to compute 95% confidence intervals.

fallacy" (*32*), i.e., mistakenly giving causal interpretation to a noncausally identified confounder that is included in a model to account for endogeneity. This could be the case if, for instance, ego-state policies are in part determined by alter state policy and/or mobility levels, or whether both ego state mobility and alter state policy are in part determined by unobserved confounding factors, such as similarities in the demographic composition of connected states. It is also possible that our spillover analysis suffers from what epidemiologists refer to as "the ecological fallacy" (*33*), as we are unable to observe and analyze human mobility outcomes for the specific subsets of people in a given ego county that are socially connected to each possible alter county.

When applied in the context of COVID-19–related mobility interventions, our analysis framework suggests that local social distancing or shelter-in-place orders may only have limited effectiveness and that the totality of the policies implemented by "peers" is approximately three times more effective at reducing mobility compared to the local policy alone. If our causal estimates are accurate, then they indicate that such policy contrasts may create scenarios in which closure policies are counterproductive (*34*, *35*), as those living in locked down regions may travel to reopened regions, potentially causing new COVID-19 hot spots to emerge. Our analysis provides suggestive evidence that such travel spillovers are not only systematic and predictable but also large and, thus, meaningful to our public health. These points are especially important to consider in light of the tiered reimposition of social distancing measures in the United Kingdom or the regional approach taken in California. Our results also suggest that the effects of COVID-19–related mobility interventions are approximately symmetric, i.e., the increase in mobility due to reopenings is roughly equal to the decrease in mobility due to closures. This analysis not only furthers our understanding of the effectiveness of past mobility-related interventions but may also prove helpful in the coming months as local governments may need to reimpose mobility restrictions in response to new COVID-19 hot spots.

Last, beyond human mobility-related interventions, analysis frameworks such as ours may be useful for studying other high-impact policy interventions enacted at the local level, in some cases with little to no coordination. The most pressing example of such policy situations related to COVID-19 is that of vaccine distribution, both within the United States and abroad. State- and country-level variations in vaccine procurement and distribution plans may affect vaccine availability for other states, countries, and regions, as well as the receptivity of a region's citizens to various distribution plans. Frameworks such as ours could also be applied to non–COVID-related policies such as firearms legislation, policing, criminal law enforcement, legislation governing the labeling and spread of online misinformation, and interventions aimed at combating climate change.

Downloaded from https://www.science.org on July 24, 2022

MATERIALS AND METHODS

Model specifications

We begin with a basic DiD model that only incorporates each county's own state policy specified as follows

$$\log\left(Y_{it}\right) = \delta_{(\mathrm{ip})}D_{it}^{(\mathrm{ip})} + \delta_{(\mathrm{sh})}D_{it}^{(\mathrm{sh})} + \delta_{(\mathrm{ro})}D_{it}^{(\mathrm{ro})} + \boldsymbol{\theta}'\mathbf{X_{it}} + \alpha_i + \tau_t + \epsilon_{it}, \quad (1)$$

where $\log(Y_{it})$ is the log-transformed mobility outcome. The policy variables, $D_{it}^{(\mathrm{ip})}$, $D_{it}^{(\mathrm{sh})}$, and $D_{it}^{(\mathrm{ro})}$ are binary indicators that take the value 1 once county i is subject to a statewide closure policy of any sort (ip), a shelter-in-place order (sh), and reopening (ro), respectively. The associated parameters $\delta_{(\mathrm{ip})}$, $\delta_{(\mathrm{sh})}$, and $\delta_{(\mathrm{ro})}$ estimate the average marginal mobility differences for the corresponding policy periods. $\mathbf{X_{it}}$ denotes a set of time-varying control variables—county-level weather, county-level unemployment, county-level morbidity and mortality, state-level unemployment, and state-level morbidity and mortality—while $\boldsymbol{\theta}'$ is a vector of the corresponding parameters. Last, α_i and τ_t denote a set of county and time fixed effects and $\mathtt{D}_{it}$ denotes the error term.

We extend this base specification to capture spillover effects with the following specification

$$\begin{aligned}\log(Y_{it}) = \; & \delta_{(\mathrm{ip})}D_{it}^{(\mathrm{ip})} + \delta_{(\mathrm{sh})}D_{it}^{(\mathrm{sh})} + \delta_{(\mathrm{ro})}D_{it}^{(\mathrm{ro})} + \gamma_{(\mathrm{ip})}D_{-it}^{(\mathrm{ip})} + \\ & \gamma_{(\mathrm{sh})}D_{-it}^{(\mathrm{sh})} + \gamma_{(\mathrm{ro})}D_{-it}^{(\mathrm{ro})} + \boldsymbol{\theta}'\mathbf{X_{it}} + \alpha_i + \tau_t + \epsilon_{it}\end{aligned} \quad (2)$$

where $D_{-it}^{(\mathrm{ip})}$, $D_{-it}^{(\mathrm{sh})}$, and $D_{-it}^{(\mathrm{ro})}$ denote the weighted average of alter states' policies, weighted by Facebook connectedness, and where the cross-state policy spillovers are captured by the terms $\gamma_{(\mathrm{ip})}$, $\gamma_{(\mathrm{sh})}$, and $\gamma_{(\mathrm{ro})}$, respectively. For these specifications, we limit our analysis to the 2583 counties with a daily mean device count of at least 500 to minimize measurement error induced by SafeGraph's differential privacy algorithm. Results presented in the main text cluster SEs at the level of the ego county U.S. state. Our results remain statistically significant when quantifying uncertainty by calculating cluster- and adjacency-robust SEs (see section S4.5).

To measure the impact of policy on cross-county mobility, we use the following specifications

$$\log\left(Y_{o\to d,t}\right) = \sum_m \lambda_m D_{ot}^m + \sum_n \psi_n D_{dt}^n + \boldsymbol{\theta_1'}\mathbf{X_{ot}} + \boldsymbol{\theta_2'}\mathbf{X_{dt}} + \alpha_{o\to d} + \tau_t + \epsilon_{o\to d,t} \quad (3)$$

$$\begin{aligned}\log\left(Y_{o\to d,t}\right) = & \sum_m \lambda_m D_{ot}^m + \sum_n \psi_n D_{dt}^n + \sum_m \sum_n \pi_{m,n}\left(D_{ot}^m * D_{dt}^n\right) + \\ & \boldsymbol{\theta_1'}\mathbf{X_{ot}} + \boldsymbol{\theta_2'}\mathbf{X_{dt}} + \alpha_{o\to d} + \tau_t + \epsilon_{o\to d,t}\end{aligned} \quad (4)$$

Here, $\log(Y_{o\to d,t})$ refers to the log-transformed cross-state mobility from an origin county o moving to a destination county in a different state d on date t. Origin and destination policies are denoted by D_{ot}^m and D_{dt}^n, respectively, where $m, n \in \{(\mathrm{ip}), (\mathrm{sh}), (\mathrm{ro})\}$; $\mathbf{X_{ot}}$ and $\mathbf{X_{dt}}$ denote time-varying origin and destination control variables covering county-level weather, county-level morbidity and mortality, and state-level morbidity and mortality for both the origin and destination. $\alpha_{o\to d}$ and τ_t correspond to directed pair and time fixed effects, and $\mathtt{D}_{o\to d,t}$ represents the error term. Equation 3 models the impacts of origin and destination policies linearly, whereas Eq. 4 includes all possible interactions between origin and destination policies.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/31/eabe7733/DC1

REFERENCES AND NOTES

1. A. G. Chandrasekhar, P. Goldsmith-Pinkham, M. O. Jackson, S. Thau, Interacting regional policies in containing a disease. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021520118 (2021).
2. R. M. Bond, C. J. Fariss, J. J. Jones, A. D. Kramer, C. Marlow, J. E. Settle, J. H. Fowler, A 61-million-person experiment in social influence and political mobilization. *Nature* **489**, 295–298 (2012).
3. S. G. Bronars, J. R. Lott Jr., Criminal deterrence, geographic spillovers, and the right to carry concealed handguns. *Am. Econ. Rev.* **88**, 475–479 (1998).
4. S. Aral, C. Nicolaides, Exercise contagion in a global social network. *Nat. Commun.* **8**, 14753 (2017).
5. S. Gupta, T. D. Nguyen, F. L. Rojas, S. Raman, B. Lee, A. Bento, K. I. Simon, C. Wing, "Tracking public and private response to the COVID-19 epidemic: Evidence from state and local government actions," *Tech. rep.* (National Bureau of Economic Research, 2020).
6. J. Sears, J. M. Villas-Boas, V. Villas-Boas, S. B. Villas-Boas, *Are we #stayinghome to flatten the curve?* (Department of Agricultural and Resource Economics, first version, 5 April 2020).
7. S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod; Imperial College COVID-19 Response Team, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S. Bhatt, Estimating the effects of non-pharmaceutical interventions on COVID-19 in europe. *Nature* **584**, 257–261 (2020).
8. S. Hsiang, D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, T. Wu, The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).
9. H. Allcott, L. Boxell, J. C. Conway, B. A. Ferguson, M. Gentzkow, B. Goldman, "What explains temporal and geographic variation in the early us coronavirus pandemic?," *Tech. rep.* (National Bureau of Economic Research, 2020).
10. D. Holtz, M. Zhao, S. G. Benzell, C. Y. Cao, M. A. Rahimian, J. Yang, J. Allen, A. Collis, A. Moehring, T. Sowrirajan, D. Ghosh, Y. Zhang, P. S. Dhillon, C. Nicolaides, D. Eckles, S. Aral, Interdependence and the cost of uncoordinated responses to COVID-19. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19837–19843 (2020).
11. L. Di Domenico, G. Pullano, C. E. Sabbatini, P.-Y. Boëlle, V. Colizza, Expected impact of reopening schools after lockdown on COVID-19 epidemic in île-de-france. medRxiv 2020.05.08.20095521 [Preprint]. 12 May 2020. https://doi.org/10.1101/2020.05.08.20095521.
12. E. L. Glaeser, G. Z. Jin, B. T. Leyden, M. Luca, "Learning from deregulation: The asymmetric impact of lockdown and reopening on risky behavior during COVID-19" (Working Paper 27650, National Bureau of Economic Research, 2020).
13. C. Xiong, S. Hu, M. Yang, W. Luo, L. Zhang, Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 27087–27089 (2020).
14. T. D. Nguyen, S. Gupta, M. Andersen, A. Bento, K. I. Simon, C. Wing, "Impacts of state reopening policy on human mobility," *Tech. rep.* (National Bureau of Economic Research, 2020).
15. S. Chang, E. Pierson, P. W. Koh, J. Gerardin, B. Redbird, D. Grusky, J. Leskovec, Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**, 82–87 (2021).
16. C. O. Buckee, S. Balsari, J. Chan, M. Crosas, F. Dominici, U. Gasser, Y. H. Grad, B. Grenfell, M. E. Halloran, M. U. Kraemer, M. Lipsitch, C. J. E. Metcalf, L. A. Meyers, T. A. Perkins, M. Santillana, S. V. Scarpino, C. Viboud, A. Wesolowski, A. Schroeder, Aggregated mobility data could help fight COVID-19. *Science* **368**, 145–146 (2020).
17. N. Oliver, B. Lepri, H. Sterly, R. Lambiotte, S. Deletaille, M. De Nadai, E. Letouzé, A. A. Salah, R. Benjamins, C. Cattuto, V. Colizza, N. de Cordes, S. P. Fraiberger, T. Koebe, S. Lehmann, J. Murillo, A. Pentland, P. N. Pham, F. Pivetta, J. Saramäki, S. V. Scarpino, M. Tizzoni, S. Verhulst, P. Vinck, Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci. Adv.* **6**, eabc0764 (2020).
18. A. L. Olsen, F. Hjorth, "Willingness to distance in the COVID-19 pandemic," *Tech. rep.* (University of Copenhagen, 2020).
19. M. Painter, T. Qiu, "Political beliefs affect compliance with COVID-19 social distancing orders," *Tech. rep.* (2020). Available at SSRN 3569098.
20. H. Allcott, L. Boxell, J. Conway, M. Gentzkow, M. Thaler, D. Yang, Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *J. Public Econ.* **191**, 104254 (2020).
21. L. Chiou, C. Tucker, "Social distancing, internet access and inequality," *Tech. Rep. 26982* (National Bureau of Economic Research, 2020).
22. A. Brzezinski, V. Kecht, D. Van Dijcke, A. L. Wright, "Belief in science influences physical distancing in response to COVID-19 lockdown policies" (University of Chicago, Becker Friedman Institute for Economics Working Paper, 2020).
23. A. Simonov, S. K. Sacher, J.-P. H. Dubé, S. Biswas, "The persuasive effect of fox news: Non-compliance with social distancing during the COVID-19 pandemic," *Tech. rep.* (National Bureau of Economic Research, 2020).
24. E. Ash, S. Galletta, D. Hangartner, Y. Margalit, M. Pinna, "The effect of fox news on health behavior during COVID-19" (2020). Available at SSRN 3636762.

Downloaded from https://www.science.org on July 24, 2022

25. M. J. Menne, I. Durre, R. S. Vose, B. E. Gleason, T. G. Houston, An overview of the global historical climatology network-daily database. *J. Atmos. Oceanic Tech.* **29**, 897–910 (2012).
26. J. Raifman, K. Nocka, D. Jones, J. Bor, S. Lipson, J. Jay, M. Cole, N. Krawczyk, P. Chan, S. Galea, Covid-19 US state policy database, V15 (2020).
27. M. Bailey, R. Cao, T. Kuchler, J. Stroebel, A. Wong, Social connectedness: Measurement, determinants, and effects. *J. Econ. Perspect.* **32**, 259–280 (2018).
28. Mobile fact sheet (2019); www.pewresearch.org/internet/fact-sheet/mobile/.
29. R. F. Squire, What about bias in the safegraph dataset? (2019); www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset.
30. M. U. G. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott; Open COVID-Data Working Group, L. du Plessis, N. R. Faria, R. Li, W. P. Hanage, J. S. Brownstein, M. Layan, A. Vespignani, H. Tian, C. Dye, O. G. Pybus, S. V. Scarpino, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
31. V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, J. Robins, Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **21**, C1–C68 (2018).
32. D. Westreich, S. Greenland, The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *Am. J. Epidemiol.* **177**, 292–298 (2013).
33. S. Piantadosi, D. P. Byar, S. B. Green, The ecological fallacy. *Am. J. Epidemiol.* **127**, 893–904 (1988).
34. J. S. Jia, X. Lu, Y. Yuan, G. Xu, J. Jia, N. A. Christakis, Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* **582**, 389–394 (2020).
35. M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini Jr., A. Vespignani, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
36. C. Dwork, A. Roth, The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.* **9**, 211–407 (2014).
37. A. Goodman-Bacon, "Difference-in-differences with variation in treatment timing," *Tech. rep.* (National Bureau of Economic Research, 2018).
38. T. G. Conley, Gmm estimation with cross sectional dependence. *J. Econom.* **92**, 1–45 (1999).
39. D. Eckles, R. F. Kizilcec, E. Bakshy, Estimating peer effects in networks with peer encouragement designs. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7316–7322 (2016).
40. T. G. Conley, F. Molinari, Spatial correlation robust inference with errors in location or distance. *J. Econom.* **140**, 76–96 (2007).
41. H. H. Kelejian, I. R. Prucha, Hac estimation in a spatial framework. *J. Econom.* **140**, 131–154 (2007).
42. A. C. Cameron, J. B. Gelbach, D. L. Miller, Robust inference with multiway clustering. *J. Bus. Econ. Stat.* **29**, 238–249 (2011).
43. R. A. Fisher, *The Design of Experiments* (Oliver and Boyd, 1935).
44. G. W. Imbens, D. B. Rubin, *Causal Inference in Statistics, Social, and Biomedical Sciences* (Cambridge Univ. Press, 2015).
45. J. G. MacKinnon, M. D. Webb, Randomization inference for difference-in-differences with few treated clusters. *J. Econom.* **218**, 435–450 (2020).
46. J. Bowers, M. M. Fredrickson, C. Panagopoulos, Reasoning about interference between units: A general framework. *Political Analysis* **21**, 97–124 (2013).
47. R. Frisch, F. V. Waugh, Partial time regressions as compared with individual trends. *Econometrica* **1**, 387–401 (1933).
48. T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (KDD, 2016), pp. 785–794.
49. R. C. Team, R: A language and environment for statistical computing (2013).
50. G. Rossum, *Python Reference Manual* (Centre for Mathematics and Computer Science, 1995).
51. W. McKinney, Data structures for statistical computing in Python, in *Proceedings of the 9th Python in Science Conference* (Austin, TX, 2010), vol. 445, pp. 51–56.
52. J. Ooms, The jsonlite package: A practical and consistent mapping between JSON data and R objects. arXiv preprint arXiv:1403.2805 [stat.CO] (12 March 2014).
53. H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, H. Yutani, Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
54. S. Gaure, lfe: Linear group fixed effects. *R Journal* **5**, 104–117 (2013).
55. T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou, M. Li, J. Xie, M. Lin, Y. Geng, Y. Li, *xgboost: Extreme Gradient Boosting*, R package version 0.90.0.1 (2019).
56. R. Analytics, domc: Foreach parallel adaptor for the multicore package, *R package version 1* (2014).
57. M. Hlavac, stargazer: Well-formatted regression and summary statistics tables, *R package version 5* (2015).
58. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).

Acknowledgments: We thank the MIT Initiative on the Digital Economy for support and SafeGraph, Facebook, and the NOAA for critical data. **Author contributions:** M.Z., D.H., and S.A. all contributed to the study design, data analysis, and writing and revising the manuscript. **Competing interests:** M.Z. declares that they have no competing interests. D.H. was previously an employee and contractor of Facebook. S.A. has recently received a grant from Facebook for other research. **Data and materials availability:** Access to SafeGraph's COVID-19 mobility data can be requested at www.safegraph.com/covid-19-data-consortium. Facebook's SCI is available on the Humanitarian Data Exchange at the following link: https://data.humdata.org/dataset/social-connectedness-index. State-level social distancing and reopening data are openly available from the COVID-19 U.S. state policy database at the following link: https://github.com/USCOVIDpolicy/COVID-19-US-State-Policy-Database. Weather data are openly available at the following link: www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-ghcn. U.S. COVID-19 case and death counts are compiled by the *New York Times* based on reports from state and local health agencies and are openly available at the following link: https://github.com/nytimes/covid-19-data. U.S. unemployment data are openly available from the U.S. Bureau of Labor Statistics at the following link: www.bls.gov/. Code is available at https://github.com/mfzhao/covid_peerfx_reopening. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 17 September 2020
Accepted 14 June 2021
Published 28 July 2021
10.1126/sciadv.abe7733

Citation: M. Zhao, D. Holtz, S. Aral, Interdependent program evaluation: Geographic and social spillovers in COVID-19 closures and reopenings in the United States. *Sci. Adv.* **7**, eabe7733 (2021).

Downloaded from https://www.science.org on July 24, 2022

ScienceAdvances

Interdependent program evaluation: Geographic and social spillovers in COVID-19 closures and reopenings in the United States

Michael ZhaoDavid HoltzSinan Aral

Sci. Adv., 7 (31), eabe7733. • DOI: 10.1126/sciadv.abe7733

View the article online
https://www.science.org/doi/10.1126/sciadv.abe7733
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org on July 24, 2022

Use of this article is subject to the Terms of service

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.
Copyright © 2021 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).